

What is claimed is:

1. A method for performing free space collection in an information storage system having storage units in which information segments are located, the method comprising the steps of:

determining a fitness value for at least some of the segments by determining the product of the amount of free space in the segment and the expected time the free space will last;

choosing segments for free space collection with the maximum fitness values.

2. A method as claimed in claim 1, wherein the expected time the free space in a given segment will last is estimated by the rate of use of still-in-use data in the segment.

3. A method as claimed in claim 1, wherein the fitness value is determined by the equation:

$$\text{Fitness} = \text{Segment age} \times \frac{\text{Free space}^2}{\text{Used space}}$$

wherein segment age is the interval between a current time and a time at which the data in the segment was last written.

4. A method as claimed in claim 3, wherein the segment age is measured in one of real time, number of write

transactions, number of segments written or destage
sequence number.

5. A method as claimed in claim 3, wherein the segment
age is one of an average, maximum or minimum of the ages
of the data in the segment.

6. A method as claimed in claim 3, wherein the segment
age also includes a write interval for the segment.

7. A method as claimed in claim 1, wherein a fitness
value is determined for one of each of a group of sorted
segments.

8. A method as claimed in claim 7, wherein the segments
are sorted by the amount of used data in each segment,
each group of segments having a given range of amount of
used data.

9. A method as claimed in claim 8, wherein all full
segments are sorted into one group and the remaining
groups have equal ranges of amounts of used data.

10. A method as claimed in claim 7, wherein the segments
are listed within each group in a first-in first-out list
and the fitness value is determined for the segment at
the head of the list.

11. A method as claimed in claim 7, wherein the segments are sorted within each group by the time at which the data in the segment was last written or by a destage sequence number and the fitness value is determined for the segment with the greatest age in each group.

12. A method as claimed in claim 11, wherein the segments are sorted by a tree data structure.

13. A method as claimed in claim 1, wherein the method is used in a log structured information storage system.

14. An information storage system comprising:

a storage controller; and

storage units in which information segments are located;

the storage controller including:

means for determining a fitness value for a number of segments by determining the product of the amount of free space in each segment and the expected time the free space will last; and

means for choosing segments with the maximum fitness values, the chosen segments being used for free space collection.

15. An information storage system as claimed in claim 14, wherein a means for determining the expected time the

free space in a given segment will last is estimated by the rate of use of still-in-use data in the segment.

16. An information storage system as claimed in claim 14, wherein the means for determining a fitness value implements the equation:

$$\text{Fitness} = \text{Segment age} \times \frac{\text{Free space}^2}{\text{Used space}}$$

wherein segment age is the interval between a current time and a time at which the data in the segment was last written.

17. An information storage system as claimed in claim 16, wherein each segment is assigned a segment age based on one of the group of real time, number of write transactions or destage sequence number.

18. An information storage system as claimed in claim 14, wherein the segments are sorted into groups and the means for determining a fitness value is applied to one segment from each group.

19. An information storage system as claimed in claim 18, wherein the segments are sorted by the amount of used data in each segment, each group of segments having a given range of amount of used data.

20. An information storage system as claimed in claim 19, wherein all full segments are sorted into one group and the remaining groups have equal ranges of amounts of used data.

5

21. An information storage system as claimed in claim 18, wherein a first-in first-out list is provided within each group to list the segments and the fitness value is determined for the segment at the head of the list.

10

22. An information storage system as claimed in claim 18, wherein the segments are sorted within each group by the time at which the data in the segment was last written or by a destage sequence number.

15

23. An information storage system as claimed in claim 22, wherein a tree data structure is provided for sorting the segments.

20

24. An information storage system as claimed in claim 14, wherein the information storage system is a log structured storage system.

25

25. An information storage system as claimed in claim 14, wherein the storage system is a log structured file system.

26. An information storage system as claimed in claim 14, wherein the information storage system is an external storage system and is a log structured array.

5 27. An information storage system comprising:
a storage controller; and

storage units in which information segments are located, the storage controller performs a free space collection process by performing the steps of:

10 determining a fitness value for at least some of the segments by determining the product of the amount of free space in the segment and the expected time the free space will last;

15 choosing segments for free space collection with the maximum fitness values.